APPENDIX E

ESTIMATES OF ENVIRONMENTAL VARIANCE FOR PCC ANALYSIS

Paul McElhany, National Marine Fisheries Service, Northwest Fisheries Science Center
John C. Payne, University of Washington
Dan Rawding, Washington Department of Fish and Wildlife
Sarah Sydor, National Marine Fisheries Service, Northwest Fisheries Science Center

Overview

Environmental variance is a key parameter in the extinction risk model used to set population change criteria (PCC). The basic approach we used to estimate environmental variance is the slope method, described in Holmes (2001). This method helps correct for the large upward bias in the variance estimate that is produced by measurement error. The basic equation of the slope method is:

$$\hat{\sigma}^2 = \text{slope of var}\left(\ln\left(\frac{N_{t+\tau}}{N_t}\right)\right) \text{vs.}\ \tau\ ,$$

where N_t is a running sum of spawner abundance counts, and τ is the temporal lag between the values used for the variance estimate. For our variance estimations, we used a running sum of four years and estimated the slope based on maximum τ of 4, as did McClure et al. (2003).

In estimating extinction risk, we need to know the natural variance, because it affects the populations no matter what human actions are taken. The presence of hatchery-origin spawners can complicate the effort to determine natural variability because changes in hatchery output can uncouple observations of spawner abundance and natural population dynamics. To correct for this potential problem, we explored modifying the equation to estimate variance when natural-origin spawners are present (McElhany and Payne in prep; McClure et al. 2003). Conceptually, the correction involves modifying the N_{t+1}/N_t ratio (Table E.1).

Harvest can also mask a population's underlying variability, but we can apply corrections similar to those made for hatcheries (Table E.1) (McElhany and Payne in prep.). Although other human activities can potentially impact variability estimates, we apply the corrections to hatcheries and harvest primarily because we have *a priori* reasons to expect them to modify the variance and because data are available. Hatchery production has varied widely in some systems, leading us to suspect it influences variance estimates of the available time series. Most harvest strategies have the goal—explicit or implicit—of reducing variability on the spawning grounds. Thus we suspect that uncorrected variance estimates tend to underestimate natural variability. Applying the corrections in age-structured salmon populations requires estimating the average age of spawner return. An important issue regarding these corrections in practice is that we seldom know the measurement error in estimates of fraction of hatchery-origin spawners and the

Table E.1 Modifications to the N_{t+1}/N_t ratio to correct for harvest and hatchery impacts on the time series.

Correct for Harvest	Correct for Hatchery	Ratio[a]
No	No	$\frac{S_{t+1}}{S_t}$
No	Yes	$\frac{W_{t+1}}{W_t + \delta H_t}$
Yes	No	$\frac{S_{t+1} + C_{t+1}}{S_t}$
Yes	Yes	$\frac{W_{t+1} + C_{t+1}}{W_t + \delta H_t}$

[a] These equations ignore the complications of age structure, which are dealt with in McElhany and Payne (in prep) and McClure et al. (2003).

Key:
S_t, = total number of spawners
W_t = number of natural-origin spawners at time t
C_t = additional number of natural-origin fish that would have returned to spawn had there been no harvest
H_t = number of hatchery-origin fish that spawn in the wild
δ = reproductive success of hatchery-origin fish spawning in the wild relative to natural-origin fish

number of additional natural-origin fish that would have returned had there been no harvest. This uncertainty about the corrections input parameters may render the uncorrected estimates more reliable, even if hatcheries and harvests both influence the spawner time series.

Population-Specific Versus Pooled Variance Estimates

Because of differences in environmental conditions, every population probably has a different mean environmental variance. If we had precise and accurate estimates of the variances, we could use these data to parameterize population-specific viability curves. However, there is often much uncertainty surrounding the variance estimate, thus more accurate viability curves may be generated by pooling variance estimates from several populations, which can be averaged to produce a “generic” viability curve that can be applied to a number of populations. The PCC targets would likely still be different for all populations because target size is a function of current size, and populations likely differ in current abundance. If populations are pooled, the assumption is that they all have a similar environmental variance and that most observed differences in individual variance result from estimation error about a common mean; further, that the differences do not reflect the true underlying population-specific difference. We suspect that, in general, differences in variance estimates do not reflect population-specific estimates because there is such a high level of uncertainty about any particular population estimate. For example, if the slope method is applied to 20 years of data, only three degrees of freedom are available for the variance estimate. This results in a high level of uncertainty about the true value of σ^2 (Figure E.1); if the point estimate is 0.05, there is roughly a 32% chance the true variance is greater than 0.1, which has a large impact on the viability curve. If populations are pooled, and it is assumed that every population represents an independent variance estimate, the point estimate becomes the average of the population variance estimates, and the degrees of freedom is the sum of the degrees of freedom from each population estimate. If populations are pooled such that there are 20 degrees of freedom, the probability that a point estimate of 0.05 comes from a sample with a true value of 0.1 drops to about 3%. The individual populations are likely to reflect independent measures of variability because populations are defined based on a high level of demographic independence. The approach we have taken thus far is to pool the estimates within an ESU to estimate environmental variance.

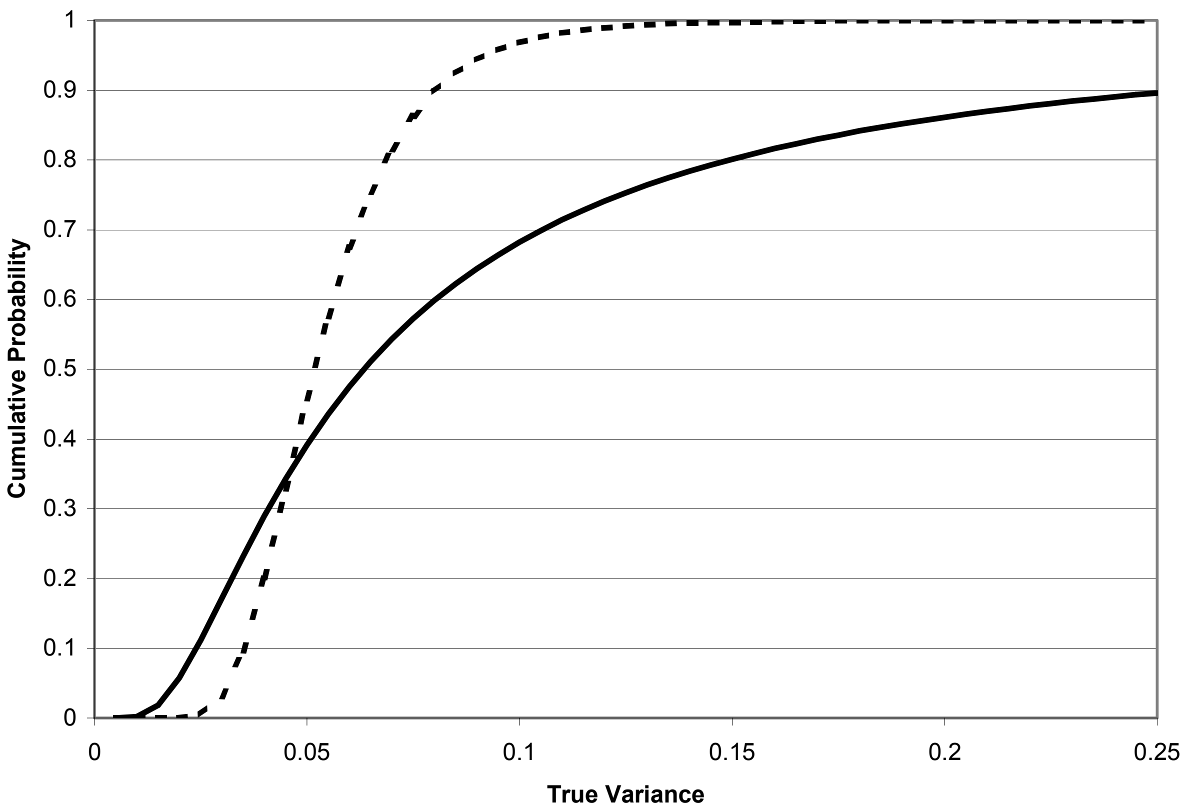

Figure E.1 Sampling distribution of variance estimate with a point estimate of 0.05. The solid line shows the distribution with 3 degrees of freedom, and the dashed line with 20 degrees of freedom.

Table E.2 Variance estimate diagnostics.

Assumption	Test
$\sigma^2 > 0$	Examine output of slope estimate
The relationship of and τ is linear	R^2 of least squared fit of variance estimate versus τ.
is normally distributed	Test for significant outliers using dffits statistics > 2
No temporal trends in	Parametric significance test and R^2 of least squared fit of versus time.
No serial autocorrelation in	Test still in development
No density dependence in time series	We have not yet conducted tests of density dependence. (These tests tend to have little power.)

Variance Estimate Diagnostics

Several diagnostic tests are available to evaluate whether the data in a time series are consistent with the assumptions of the basic demographic model. We explored tests that are similar (but not identical) to those of McClure et al. (2003) (Table E.2).

Variance Estimates for WLC Populations

The primary data needed to calculate the variance are time series of population spawner counts or of an index that is proportional to the population spawner counts. To apply the hatchery and harvest corrections we also need estimates of the fraction of hatchery-origin spawners present each year, the relative reproductive success of the hatchery-origin spawners, the number of additional natural-origin fish that would have returned had there been no harvest each year, and an estimate of the average age at spawning. We have obtained as many relevant time series as possible for populations in the Willamette/Lower Columbia domain. These time series, their references, and dataset descriptions are available on the Web at http://research.nwfsc.noaa.gov/cbd/trt/wlc_trt/viability_report.htm. A computer program that calculates the variance estimates with user provide inputs (including options for the harvest and hatchery corrections), SimSalmon version 4.5.3 beta, is available at the same Web site.

In estimating variance for WLC populations, we were limited to a large extent by available data. We explored the variance estimates under a number of different assumption options (Table E.3). The input data were collected using a variety of methods and are of mixed quality. The variance estimates and diagnostic outputs for the WLC populations under one set of options are shown in Table E.4. The variance estimates and diagnostics for all populations under all assumption option sets are available on the Web at http://research.nwfsc.noaa.gov/cbd/trt/wlc_trt/viability_report.htm. The average variance estimate by ESU and life-

Table E.3 Analysis options for estimating environmental variance from available time series in the WLC.[a]

Option Number	Relative Fitness of Hatchery Origin Spawners[b]	Includes Correction for Harvest?[c]	Years Used for Analysis[d]
1	0	No	All data
2	0.5	No	All data
3	1	No	All data
4	0	Yes	All data
5	0.5	Yes	All data
6	1	Yes	All data
7	0	No	Since 1980
8	0.5	No	Since 1980
9	1	No	Since 1980
10	0	Yes	Since 1980
11	0.5	Yes	Since 1980
12	1	Yes	Since 1980

[a] The hatchery correction was applied to all options. Because of limited data availability or the history of the population, not all populations could be analyzed under all the options.
[b] The relative reproductive success of hatchery-origin spawners compared to natural origin spawners assumed for a particular option.
[c] Indicates whether or not the harvest correction was applied for a particular option.
[d] Indicates whether or not the analysis used all available data or only data since 1980 for a particular option.

history type for each option is shown in Table E.5. Table E.6 shows the summed degrees of freedom associated with the averages in Table E.5. For a given ESU/life-history type, the variance averages are relatively similar under all assumption option sets.

Table E.4 Variance estimates and diagnostics for WLC populations, assuming that hatchery fish have the same reproductive success as natural-origin fish and with no harvest correction.[a]

ESU	Population	Years of Data	Sample Size for (N_{t+1}/N_t)	Variance (95% Confidence Interval)	Variance Degrees of Freedom	Slope of (N_{t+1}/N_t) vs. Time	Number of Outliers from Normal Distribution
Lower Columbia chinook spring	Cowlitz River	1980–1999	16	0.015 (0.005–0.164)	3.37	n.s.[b]	1
Lower Columbia chinook salmon late fall	Lewis River	1964–2000	17	0.038 (0.017–0.152)	7.37	n.s.	1
	Sandy River	1984–2001	9	0.04 (0.013–0.591)	2.9	-0.043	0
Lower Columbia chinook fall	Big White Salmon River	1964–2000	17	0.175 (0.078–0.691)	7.37	n.s.	1
	Coweeman River	1964–2000	17	0.186 (0.083–0.735)	7.37	n.s.	1
	Cowlitz River	1964–2000	17	0.714 (0.317–2.817)	7.37	n.s.	0
	East Fork Lewis River	1980–2000	17	0.01 (0.003–0.094)	3.61	n.s.	0
	Elochoman River	1964–2000	17	0.381 (0.169–1.505)	7.37	n.s.	1
	Grays River	1964–2000	17	0.31 (0.138–1.224)	7.37	n.s.	1
	Kalama River	1964–2000	17	0.311 (0.138–1.226)	7.37	n.s.	1
	Mill Creek River	1980–2000	17	0.141 (0.049–1.382)	3.61	-0.028	0
	Washougal River	1964–2000	17	0.088 (0.039–0.346)	7.37	n.s.	0
	Wind River	1980–2000	12	0.361 (0.125–3.534)	3.61	n.s.	0
	Clackamas River	1967–2001	26	0.091 (0.04–0.384)	6.9	n.s.	2
Lower Columbia steelhead winter	Clackamas River	1958–2001	40	0.097 (0.046–0.321)	9.02	n.s.	2
	Kalama River	1977–2002	22	0.031 (0.012–0.197)	4.78	n.s.	2
	North Fork Toutle River	1989–2002	10	0.001 (0–0.053)	1.97	n.s.	1
	South Fork Toutle River	1984–2002	10	0 (0–0.002)	3.14	n.s.	2
	Sandy River	1978–2001	16	0.027 (0.01–0.202)	4.31	n.s.	0
	East Fork Lewis River	1985–1994	6	0.004 (0.001–3.798)	1.02	n.s.	0
	Hood River	1992–2000	5	0.041 (no estimate)	0.79	0.133	1

Lower Columbia steelhead summer	Kalama River	1977–2003	23	0.178 (0.07–1.068)	5.02	n.s.	2
	Washougal River	1986–2003	14	0.07 (0.022–1.049)	2.9	0.03	2
	Wind River	1989–2003	11	0.006 (0.002–0.182)	2.2	n.s.	1
	Hood River	1992–2000	5	0.01 (no estimate)	0.79	0.104	0
Lower Columbia chum	Grays River	1967–2000	28	0.051 (0.022–0.222)	6.67	n.s.	1
	Hardy Creek	1957–2000	40	0.076 (0.036–0.253)	9.02	n.s.	2
	Lower gorge	1944–2000	53	0.08 (0.041–0.216)	12.07	n.s.	3
Upper Willamette chinook salmon spring	Clackamas River	1958–2002	41	0.107 (0.051–0.348)	9.25	n.s.	3
	McKenzie River	1970–2001	28	0.122 (0.051–0.572)	6.19	n.s.	1
Upper Willamette steelhead winter	Calapooia River	1980–1997	14	0.211 (0.067–3.147)	2.9	n.s.	1
	Molalla River	1980–1997	14	0.072 (0.023–1.068)	2.9	n.s.	0
	North Santiam River	1980–1997	14	0.066 (0.021–0.984)	2.9	n.s.	1
	South Santiam River	1980–1997	14	0.008 (0.002–0.113)	2.9	n.s.	0

[a] Option 3 in Table E.3.
[b] n.s. indicates that slope is not significant at $\alpha = 0.05$).

Table E.5 Average variance estimates for WLC ESU/life-history types under a number of different assumption sets.[a]

Option Number	Lower Columbia						Upper Willlamette	
	Chinook Salmon			Steelhead		Chum	Chinook	Steelhead
	Spring	Late Fall	Fall	Winter	Summer		Spring	Winter
1	0.015	0.039	0.251	0.030	0.063			0.095
2	0.015	0.039	0.261	0.029	0.065			0.089
3	0.015	0.039	0.252	0.029	0.066	0.069	0.114	0.089
4		0.081	0.288	0.027	0.069			
5		0.080	0.287	0.027	0.064			
6		0.080	0.287	0.028	0.064			
7	0.015	0.039	0.251	0.029	0.064			0.095
8	0.015	0.039	0.261	0.029	0.065			0.089
9	0.015	0.039	0.268	0.029	0.065			0.089
10		0.081	0.288	0.023	0.064			
11		0.080	0.287	0.022	0.062			
12		0.080	0.287	0.022	0.062			
Total Average	0.015	0.053	0.271	0.027	0.065	0.069	0.114	0.091

[a] Table E.3 regarding assumption options. Because of data availability, some ESU/life-history types could not be evaluated under some assumption options.

Table E.6 Summed degrees of freedom estimates for WLC ESU/life history types under a number of different assumption sets.[a]

Option Number	Lower Columbia						Upper Willamette	
	Chinook Salmon			Steelhead		Chum	Chinook	Steelhead
	Spring	Late Fall	Fall	Winter	Summer		Spring	Winter
1	3	10	62	25	11			12
2	3	10	62	25	11			12
3	3	10	69	25	11	28	15	12
4		7	59	23	10			
5		7	59	23	10			
6		7	59	23	10			
7	3	7	36	19	10			12
8	3	7	36	19	10			12
9	3	7	36	19	10			12
10		4	32	17	9			
11		4	32	17	9			
12		4	32	17	9			

[a] See Table E.3 regarding assumption options. Because of data availability, some ESU/life-history types could not be evaluated under some assumption options. These are the summed degrees of freedom that accompany the variance averages in Table E.5.

Preliminary Variance Estimate Conclusion

The average variance estimates by ESU and life-history type ranged from 0.015 for Lower Columbia River spring chinook salmon to 0.287 for Lower Columbia fall chinook (Table E.5). The lowest single population variance estimate was the Wind River winter steelhead, at 0.006; the highest was Cowlitz fall chinook, at 0.714 (Table E.4). The average of the ESU/life-history averages is approximately 0.08. The Lower Columbia fall chinook had consistently higher variance estimates than other ESU/life-history types. This may reflect some inherently higher variability in the Lower Columbia fall chinook populations; alternatively, it may reflect high levels of measurement error in the abundance time series. The Lower Columbia fall chinook populations tend to have large fractions of hatchery-origin spawners, but the Washington Department of Fish and Wildlife (WDFW) considers actual hatchery fraction estimates to be very imprecise. It is interesting to note that the Lower Columbia population with little hatchery input (Coweeman) has a variance estimate of 0.187, which is below the average for Lower Columbia fall chinook but above the average for other ESU/life-history types.

Based on examination of the data and diagnostic output, we tentatively applied a variance estimate of 0.05 for all ESU populations in the WLC domain and assumed 20 degrees of freedom. This is not based a single mathematical calculation but on a professional judgment evaluation that incorporated the estimated average variances and assessment of overall data quality of individual time series. This assessment led to a discounting of the variance estimates from the Lower Columbia fall chinook population for reasons discussed in the previous paragraph. The variance estimate of 0.05 is just an initial starting point; its accuracy, used in conjunction with the PCC targets, would be expected to improve with additional high-quality time-series data.

References

Holmes, E. 2001. Estimating risks in declining populations with poor data. Proc. Nat. Acad. Sci. 98(9): 5072–5077.

McClure, M. M., Elizabeth E. Holmes, Beth L. Sanderson, and Chris Jordan. 2003. A large-scale, multi-species status assessment: Anadromous salmonids in the Columbia River Basin. Ecol. App. in press

McElhany, P., and J. Payne. 2002. SimSalmon. Computer program. Available on the Web at http://research.nwfsc.noaa.gov/cbd/trt/wlc_trt/viability_report.htm.

McElhany, P., and J. Payne. In prep. Setting viability criteria for Pacific salmonids using abundance data.